

**RATIONAL DESIGN AND SYNTHESIS OF
INHIBITORS FOR H1N1 NEURAMINIDASE AND
DENGUE PROTEASE ENZYMES**

by

MAYWAN HARIONO

**Thesis submitted in fulfillment of the
requirements for the degree of
Doctor of Philosophy**

July 2015

ACKNOWLEDGEMENT

Formost, a great thankful was expressed to The Almighty, Allah SWT. Without His Blessings, it was impossible to complete my PhD thesis in an exact time. Secondly, I would like to express my sincere gratitude to my advisor Prof. Dr. Habibah A. Wahab for her continuous support on my PhD study and research, for her patience, motivation, enthusiasm, and immense knowledge. Her guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my PhD study. A great thank was also expressed to Universiti Sains Malaysia for awarding me the USM fellowship and Malaysian Ministry of Science and Technology Inovation for fully funding my projects under ScienceFund Research Grant. Special thanks were expressed to Dr. Tan Mei Lan, Prof. Dr. Hasnah Osman, Dr. Ezatul E. Kamarulzaman for their helps during my study. I would also like thank to all laboratory technicians in Pharmaceutical Technology Dept. (in particular, En. Shamsudin and En. Rosli) and Pharmaceutical Chemistry Dept. (En. Hammid, En. Fizal, En. Zaenudin and En. Annuar) for their helps. I thank my fellow labmates: Azizairol, Dr. Belal Al-Najjar, Dr. Muchtaridi, Sufian, Yusuf and Khairi for stimulating discussion as well as our brotherhood in happy and sad moments. Dya, Neny, Rina, Nadia, Saira, Hanim, Ban Hong, Lim, Lee, Stella, Vincent, Adila, Vanee, Nasuha, Adiba, Shakina, Fakhrul, Faizul, Emah and Wani, thanks for all the funs despite the hard times. Thank you and thank you again.

TABLE OF CONTENTS

ACKNOWLEDGEMENT	ii
TABLE OF CONTENTS	iii
LIST OF TABLES	vi
LIST OF FIGURES	viii
LIST OF SCHEMES	xv
LIST OF SYMBOLS	xvii
LIST OF ABBREVIATIONS	xviii
ABSTRAK	xx
ABSTRACT	xxii
 CHAPTER ONE INTRODUCTION	 1
1.1 Statement of the Problem	1
1.2 Objectives	9
CHAPTER TWO LITERATURE REVIEW	11
2.1 Influenza A and Its Drug Treatment	11
2.1.1 Influenza A virus	11
2.1.2 Influenza A neuraminidase	14
2.1.3 Neuraminidase Inhibitors	18
2.2 Dengue and Potential Anti-dengue Compounds	25
2.2.1 Dengue virus	26
2.2.2 Non Structural Protein 3 (NS3) Protease	30
2.2.3 NS3 Protease Inhibitors	33
2.3 Computer-Aided Drug Design	43
2.3.1 Pharmacophore Modelling	44
2.3.2 Molecular Docking	45
2.3.3 Quantitative Structure-Activity Relationship (QSAR) Modelling	48
2.4 Content of the Thesis	51
CHAPTER THREE RATIONAL DESIGN AND SYNTHESIS OF H1N1 NA INHIBITORS	52

3.1	Overview	52
3.1.1	Pharmacophore Modelling	52
3.1.2	QSARs	54
3.1.3	Molecular Docking	56
3.2	Methodology	59
3.2.1	Materials	59
3.2.1 (a)	Softwares and Hardwares	59
3.2.1 (b)	Reagents and Other Consumable Materials	59
3.2.1 (c)	Instruments	60
3.2.2	Methods	61
3.2.2 (a)	Molecular Modelling	61
3.2.2 (b)	Chemical Synthesis of Ferulic Acid and Its Derivatives	64
3.2.2 (c)	H1N1 Neuraminidase Assay	81
3.2.2 (d)	QSARs	82
3.3	Results and Discussions	85
3.3.1	Pharmacophore Mapping	85
3.3.2	Molecular Docking	88
3.3.3	Synthesis of Ferulic Acid and Its Derivatives	93
3.3.4	H1N1 Neuraminidase Assay	120
3.3.5	QSARs	125
3.3.6	Design of Novel H1N1 NA Inhibitors	133
3.4	Conclusion	135
	CHAPTER FOUR RATIONAL DESIGN AND SYNTHESIS DENV2 NS2B/NS3 PROTEASE INHIBITORS	137
4.1	Overview	137
4.1.1	Pharmacophore Modelling	137
4.1.2	QSARs	139
4.1.3	Molecular Docking	139
4.2	Methodology	141
4.2.1	Materials	141
4.2.1 (a)	Softwares and Hardwares	141
4.2.1 (b)	Reagents and Other Consumable Materials	141

4.2.1 (c) Instruments	143
4.2.2 Methods	143
4.2.2 (a) Molecular Modelling	143
4.2.2 (b) Chemical Synthesis of Thioguanine Derivatives	145
4.2.2 (c) DENV2 NS2B-NS3 Protease Assay	166
4.2.2 (d) QSARs	167
4.3 Results and Discussions	170
4.3.1 Pharmacophore Mapping	170
4.3.2 Molecular Docking	176
4.3.3 Synthesis of Thioguanine Derivatives	179
4.3.4 DENV2 NS2B-NS3 Protease Assay	200
4.3.5 QSARs	209
4.3.6 Design of Novel H1N1 NA Inhibitors	227
4.3.7 Synthesis and DENV2 NS2B-NS3 Protease Inhibition Assay of the new designed compounds	229
4.4 Conclusion	232
CHAPTER FIVE FURTHER DISCUSSION	234
CHAPTER SIX CONCLUSION AND FUTURE WORK	242
5.1 Conclusion	242
5.2 Future Work	243
REFERENCES	245
APPENDICES	
Appendix 1 The published synthesised compounds and their references	
Appendix 2 Examples of Assay Data Management	
Appendix 3 The Compound's Structures Used as Test Sets	
Appendix 4 Predicted Chemical Shift of ¹ H-NMR on Ferulic Acid and Thioguanine Derivatives	
Appendix 5 Fourier Transform Infra Red (FTIR) Correlation Table	
Appendix 6 Spectroscopic Data of Ferulic Acid and Thioguanine Derivatives	
Appendix 7 Drug-Dose Response Curves of Ferulic Acid and Thioguanine Derivatives	
Appendix 8 ORTEP drawings of <i>MH005</i> and <i>MH010</i>	

LIST OF TABLES

	Page
3.1 The series of ferulic acid derivatives designed in this study	62
3.2 The list of published NA inhibitors taken for the test set.	83
3.3 The fit values of the ligands and their corresponding pharmacophore features in Model A-5-5 for NA inhibitor.	85
3.4 The ligands and their FEB, amino acid residues-ligand H-bond interaction as well as its distances produced by docking against H1N1 NA.	91
3.5 The training set compounds with their IC ₅₀ values against H1N1 NA.	126
3.6 Superimposition and the molecular properties of compounds against H1N1 NA in the training set generated using Discovery Studio 2.5.	127
3.7 The selected ligand from the test set which showed the best correlation between experimental and predicted pIC ₅₀ .	131
3.8 The designed compounds for NA inhibitors based on the QSAR modelling	134
3.9 The predicted IC ₅₀ and the descriptors of the designed compound for NA inhibitors.	134
4.1 The compounds' series of thioguanine derivatives	144
4.2 List of published DENV2 NS2B-NS3pro inhibitors taken as the test set for the QSAR study	168
4.3 Fit values of the ligands and their corresponding pharmacophore features for dengue protease inhibitor	171
4.4 The ligands and their corresponding pose, FEB, amino acid residue-ligand H-bond interaction as well as its distances produced by docking against DENV2 NS2B-NS3pro model, JMR_977_sm	177
4.5 The training set compounds with their IC ₅₀ values for dengue protease inhibitors.	210
4.6 The superimposition and the molecular properties of the training	211

	set compounds for dengue protease inhibitors generated using Discovery Studio 2.5	
4.7	The best QSAR models for dengue protease inhibitors and their regression statistics generated using Genetic Function Approximation algorithm embedded in Discovery Studio 2.5	214
4.8	The fingerprints features which is used in linear equation	218
4.9	The fingerprints features which is used in binary interaction equation	219
4.10	The fingerprints features which is used in simple cubic equation	219
4.11	The fingerprints features which is used in full cubic equation	220
4.12	The fingerprints features which is used in simple quadratic equation	221
4.13	The fingerprints features which is used in full quadratic equation	222
4.14	The superimposition and the molecular properties of the test set compounds for dengue protease inhibitors included their predicted as well as experimental IC_{50} values generated using Discovery Studio 2.5	225
4.15	The regression statistics of the test set experimental pIC_{50} versus its predicted pIC_{50} based on the QSAR modelling for dengue protease inhibitors	227
4.16	The list of designed compound based on the QSAR modelling for dengue protease inhibitors	228
4.17	The prediction of molecular properties and predicted IC_{50} values of the new designed compounds for dengue protease inhibitors	228

LIST OF FIGURES

		Page
1.1	The death number of pandemic H1N1 influenza 2009.	1
1.2	The dengue case statistic in Malaysia from 1995-2013	5
1.3	The structure of ferulic acid and vanillin	8
1.4	The structure of diversity0713 and thioguanine	9
2.1	The influenza viral lipid envelope with a nucleocapsid containing three surface proteins: hemagglutinin (HA), neuraminidase (NA), and the M2 proton ion channel. The genome of the viral RNAs are presented as red coils bound to Ribonuclear Proteins (RNPs)	13
2.2	The tetramer shapes of neuraminidase N2	15
2.3	The comparison of crystal structures between N1 and N9. (A) A superposition of N1 (PDBID 2HTY) and N9 enzyme (PDBID 1F8B) where the key residues are shown in ball-and-stick forms. (B) Ligand in the active cavity of 2HTY with adjacent 150-loop. (C) Ligand in the active cavity of 1F8B	16
2.4	The cleavage of the new synthesised virus from its sialic acid receptor by neuraminidase	17
2.5	The binding site of A/Tokyo/3/67 (H2N2) influenza virus NA and sialic acid (PDBID 2BAT). The green dot lines describe the hydrogen bonding between ligand and specific amino acid residues of NA's active site	18
2.6	'Airplane' model of NA active site	19
2.7	The structure of NA inhibitors from sialic acid derivatives	20
2.8	The structure of 4 with the dot boxes describes the nonpolar site of the particular structure	21
2.9	The X-ray crystal structure of neuraminidase complex with 4 . The hydrophobic interactions are indicated by the closeness between isopentyl group and amino acid residues Ile222,	22

	Arg224 and Ala246	
2.10	The structure of 5 in 2D as well as its 3D structure complex to N8 NA (PDBID 2HTU) was generated using Discovery Studio 2.5	24
2.11	The 3D structure of active laninamivir (6) in complex with pO9N1 NA (PDBID 2HTU)	25
2.12	The dengue virus genome organization and its cleavage processing scheme	27
2.13	The life cycle of dengue virus	30
2.14	The crystal structure of DENV2 NS2B-NS3pro (PDBID 2FOM); (a) a ribbon form and (b) a surface form was created using Discovery Studio 2.5	31
2.15	The crystal structure of DENV4 NS2B-NS3 (PDBID 2VBC) was created using Discovery Studio2.5	33
2.16	The structure of decapeptide substrate (7) for dengue NS2B-NS3pro substrate	34
2.17	The crystal structure of NS2B-NS3pro in the absence and presence of an inhibitor. (a) DENV2 NS2B-NS3pro (NS3 = gray ribbon; NS2B = yellow) (b) WNV NS2B-NS3pro in complex with Bz-Nle-Lys-Arg-Arg- H (orange)	36
2.18	The WNV NS2B-NS3pro crystal structure (PDBID 2FP7). The hydrogen bond (dotted line) interaction between substrate-based inhibitor with the residues in S1 pocket and S2 2 of NS2B-NS3	36
2.19	The structure of a tetrapeptide DENV2 NS3pro inhibitor	39
2.20	The structure of 10	40
2.21	The structure of 11	41
2.22	The structure of anthracene-based DENV2 protease inhibitor	41
2.23	The structure of 14	42
2.24	The structure of cyclohexenyl derivatives	43
3.1	The pharmacophore model GR-210729 for NA inhibitor. The pharmacophore features are color-coded: blue (NI), light yellow (HBD), PI (white), HBA (green) and H (light blue)	53

3.2	Mapping Model A-5-5 against sialic acid derivative compared with its corresponding docked poses into NA	54
3.3	The structure of 3 as a potent neuraminidase inhibitor and its 2D interaction with NA	57
3.4	The docking pose of 2D-interaction of AV5027 with NA (PDB code 3TI6)	58
3.5	Mapping pharmacophore models A-5-5 against 020 . The green-vectored spheres encoded for HBA while the blue sphere represents negative ionizable was visualized using Discovery Studio 2.5	86
3.6	Mapping pharmacophore models A-5-5 against (a) 013 , (b) 014 , (c) 015 , (d) 016 , (e) 021 and (f) 019 was visualized using Discovery Studio 2.5	87
3.7	Mapping pharmacophore models A-5-5 against (a) 000 , (b) 004 , (c) 005 , (d) 006 and (e) 007 was visualized using Discovery Studio 2.5	88
3.8	The control docking pose of 4 to H1N1 NA (PDBID 3TI6) in ribbon form (in set) and surface form was visualized using Discovery Studio 2.5	89
3.9	The proton NMR spectrum of compound 001	96
3.10	The proton NMR spectrum of compound 000	97
3.11	The mass spectrum of compound 001 is calculated using QTOF-MS as $[M+NH_4]^+$ found m/z 257.2999	98
3.12	The proton NMR spectrum of compound 002	100
3.13	The mass spectrum of compound 002 is calculated using QTOF-MS as $[M+Na]^+$ found m/z 262.1653	101
3.14	The proton NMR spectrum of compound 003	102
3.15	The FTIR spectra of compound 003 showed the presence of amino group as a broad absorption at 3420 and 3215 m^{-1} described the success of nitro reduction from compound 001	103
3.16	The mass spectrum of compound 003 is calculated using QTOF-MS as $[M+NH_4]^+$ found m/z 227.2268	103
3.17	The proton NMR spectrum of compound 004	106

3.18	The proton NMR spectrum of compound 009	107
3.19	The proton NMR spectrum of compound 012	110
3.20	The proton NMR spectrum of compound 013	113
3.21	The proton NMR spectrum of compound 019	116
3.22	The mass spectrum of compound 019 is calculated using QTOF-MS as $[M]^+$ 209.2019, found 209.1438	117
3.23	The proton NMR spectrum of compound 020	119
3.24	The drug-dose response curve of 000 against H1N1 NA	121
3.25	The docking pose of (a) 000 and (b) 018 was visualized using Discovery Studio 3.5	121
3.26	The drug dose-response curve of 001 and 002 against H1N1 NA	122
3.27	The docking pose of (a) 001 and (b) 002 to H1N1 NA (PDBID 3TI6) was visualized using Discovery Studio 2.5	123
3.28	The drug dose-response curve of 003 , 004 and 012 against H1N1 NA	124
3.29	The drug dose-response curve of 019 , 020 and 021 against H1N1 NA	125
3.30	The predicted IC_{50} versus experimental IC_{50} of the of NA inhibitors training set generated using multiple linear regression method	130
3.31	The graph plotting the experimental of pIC_{50} versus the predicted pIC_{50} based on QSAR modelling. The graph was generated using Microsoft Excell 2007	131
3.32	The similar pose of 020 produced by docking (left) and pharmacophore mapping (right) was generated using Discovery Studio 2.5	132
3.33	Proposed new model of H1N1 Neuraminidase inhibitor based on QSAR model above	133
3.34	Intermolecular interactions between (a) 022(M) and (b) 023(M) was visualized using Discovery Studio 2.5	135
4.1	Pharmacophore ligand model for dengue protease generated from HIV protease inhibitors	138

4.2	Dynamic pharmacophore model of dengue protease inhibitor	138
4.3	The docking pose of three 4-hydroxypanduratin derivatives (a) 246DA, (b) 20H46DA and (c) 2446DA against DENV2 NS2B-NS3pro (PDBID 2FOM)	140
4.4	The docking pose of RKR into the 2FOM model	141
4.5	Mapping pharmacophore models S5T5HO6 against MH019 (a) 3D structure and (b) 2D structure.	172
4.6	Mapping pharmacophore models S5T5HO6 against (a) MH010 , (b) MH011 and (c) MH020 was visualized using Discovery Studio 2.5	173
4.7	Mapping pharmacophore models S5T5HO6 against (a) MH001 , (b) MH002 and (c) MH003 was visualized using Discovery Studio 2.5	174
4.8	Mapping pharmacophore models S5T5HO6 against (a) MH007 , (b) MH008 and (c) MH009 was visualized using Discovery Studio Client 2.5	175
4.9	Mapping pharmacophore models S5T5HO6 against MH016 and MH018 was visualized using Discovery Studio 2.5.	176
4.10	The overlay of an initial pose and a control docking of tetrapeptide inhibitor (NDL1001) to the DENV2 NS2B-NS3pro from the Wichapong model was visualized using Discovery Studio 2.5	177
4.11	The proton NMR spectrum of compound MH001	180
4.12	The carbon 13 NMR spectrum of compound MH001	181
4.13	The FTIR spectrum of MH001 shows the presence of $-NH_2$ indicated the absence of alkylation at this particular functional group	183
4.14	The Mass spectrum of compound 003 is calculated using QTOF-MS as $[M+H]^+$ found m/z 224.1142	183
4.15	The proton NMR spectrum of compound MH013	187
4.16	The carbon 13 NMR spectrum of compound MH013	188
4.17	The FTIR spectrum of MH013	189
4.18	The FTIR spectrum of MH000	190

4.19	The mass spectrum of compound MH013 is calculated using QTOF-MS as $[M+H]^+$ found m/z 274.2964	190
4.20	The proton NMR spectrum of compound MH019	193
4.21	The carbon 13 NMR spectrum of compound MH019	194
4.22	The mass spectrum of MH019 is calculated using QTOF-MS as $[M]^+$ found m/z 475.2863	195
4.23	The proton NMR spectrum of compound MH022	198
4.24	The carbon-13 NMR spectrum of compound MH022	199
4.25	The FTIR spectra of MH022 showed the presence of carbonyl group as a stretching absorption at 1552cm^{-1} described the success of acylation from 6-thioguanine	200
4.26	7-Amino-4-methylcoumarin (AMC) standard curves. The AMC concentrations were prepared within 0 to 10 μM	201
4.27	Graph of DENV2 NS2B-NS3pro activity assay (protease optimum assay)	202
4.28	The Lineweaver-burk of DENV2 NS2B-NS3pro activity assay (protease optimum assay)	203
4.29	Graph of DENV2 NS2B-NS3pro activity assay (substrate optimum assay)	204
4.30	The Lineweaver-burk of DENV2 NS2B-NS3pro activity assay (substrate optimum assay).	204
4.31	The drug dose-response curve of MH000 against DENV2 NS2B-NS3pro	205
4.32	The drug-dose response curves of MH013 , MH015 and MH018 against DENV2 NS2B-NS3pro	206
4.33	The similar binding mode of (a) MH018 and (b) MH022 produced by molecular docking. The picture was generated using Discovery Studio 2.5.	207
4.34	The drug dose-response curve of MH019 and MH020 against DENV2 NS2B-NS3pro	208
4.35	The similar pose of MH019 produced by docking (left) and pharmacophore mapping (right) was generated using Discovery Studio 2.5	209

4.36	The graph plotting the experimental versus predicted pIC ₅₀ of dengue protease inhibitors was generated using GFA embedded in Discovery Studio 2.5	215
4.37	Design of the model of DENV2 NS2B-NS3pro inhibitor based on the linear model	227
4.38	The graph plots log concentration versus % inhibition against DENV2 NS2B- NS3 protease (a) MH006 with the IC ₅₀ = 2042 μM (r ² = 0.9431) and (b) MH012 with IC ₅₀ = 55 μM (r ² = 0.9388)	229
4.39	The docking pose MH012 to DENV2 NS2B-NS3pro with the protease in surface form was visualized using Discovery Studio 2.5	230
4.40	The graph plots log concentration versus % inhibition against DENV2 NS2B- NS3 protease (a) MH021 with the IC ₅₀ = 55 μM (r ² = 0.9388) and (b) MH024 with IC ₅₀ = 132 μM (r ² = 0.9096)	231
4.41	The docking pose of (a) MH022 and (b) MH024 to DENV2 NS2B-NS3pro was visualized using Discovery Studio 2.5	232
5.1	The superposition of 019 and DANA at the binding site of H1N1 NA (3TI6.pdb).	236
5.2	The superposition of MH018 and Panduratin A at the binding site of DV2 NS2B/NS3pro.	241

LIST OF SCHEMES

	Page
3.1 The scheme of the synthesis of ferulic acid derivatives	65
3.2 The scheme of the synthesis of vanillin derivatives	66
3.3 The reaction of ferulic acid (compound 000) with fuming nitric acid – glacial acetic acid to produce compound 001	93
3.4 The reaction mechanism of <i>o</i> -nitration of 000 to yield 001	95
3.5 The reaction of compound 000 with methanol using H ₂ SO ₄ as the catalyst to produce compound 004	104
3.6 The esterification mechanism of 000 to yield 004	105
3.7 The reaction of compound 009 with isopropyl bromide, TBAI as the solid phase transfer catalyst, Na ₂ CO ₃ as the base catalyst and DMF as the solvent to produce compound 012	108
3.8 The reaction mechanism of alkylation of compound 009 to yield compound 012	109
3.9 The reaction of compound 000 with 4-fluorobenzenesulfonyl chloride using pyridine as the nucleophilic catalyst to produce compound 013	111
3.10 The reaction mechanism of benzenesulfonylation of compound 000 to yield compound 013	112
3.11 The reaction mechanism of guanidine introduction of intermediate 2 to yield 019	115
3.12 The reaction of intermediate 3 with 2-ethanolamine chloride using DIPEA as the base catalyst and dichloromethane (DCM) as a solvent to produce compound 020	117
4.1 The scheme of the synthesis of thioguanine derivatives for MH001-12 (004 was declined)	146
4.2 The scheme of the synthesis of thioguanine derivatives for MH013-21	147
4.3 The scheme of the synthesis of thioguanine derivatives for MH022-24	148

4.4	The reaction step of <i>S</i> -alkylation consists of 1) initial state, 2) thioanion state and 3) the <i>S</i> -alkylated of thioguanine	179
4.5	The reaction scheme of Schiff base formation from 6-thioguanine using sodium hydroxide-ethanol as solvent	184
4.6	The reaction mechanism of Schiff base formation to yield compound <i>MH013</i>	186
4.7	The reaction scheme of benzenesulfanylation of 6-thioguanine using pyridine as catalyst	191
4.8	The reaction scheme of acylation of 6-thioguanine using either glacial acetic acid as solvent	195
4.9	The reaction mechanism of thioguanine acetylation to yield <i>MH022</i>	197

LIST OF SYMBOLS

\AA	Angstrom
Fit	Fit value
Σ	Sum
ΔG	The Free Energy
ΔS	The entropy
r	Correlation coefficient
K_i	Inhibition Constant
K_m	Michaelis Menten Constant
nM	nanoMolar
μM	microMolar
R	<i>Rectus</i> configuration
S	<i>Sinister</i> configuration
evals	Evaluation
[S]	Substrate concentration
J	Coupling constant
s	Singlet
d	Doublet
t	Triplet
q	Quartet
m	Multiplet
δ_C	Chemical shift of carbon-13
$1/V$	The reciprocal of reaction assay velocity
[E]	Enzyme concentration
r^2_{adj}	Adjusted quadratic correlation coefficient
r^2_{pred}	Predicted quadratic correlation coefficient
F	Variance
δ^+	Positive partial
δ^-	Negative partial
AlogP	Partition coefficient based on Ghose and Crippen's method

LIST OF ABBREVIATIONS

act	Actual
AMC	Aminomethyl coumarine
<i>br</i>	Broad
Boc	Butoxycarbonyl
Bz	Benzoil
C	Capsid
DANA	2-deoxy-2,3-didehydro- <i>N</i> -acetylneuraminic acid
DIPEA	<i>N, N</i> -diisopropylethylamine
E	Envelope
ECFP	Extended Connectivity Finger Print
EDG	Electron Donating Group
ER	Endoplasmic Reticulum
EWG	Electron Withdrawing Group
GFA	Genetic Function Approximation
GTP	Guanine Triphosphate
GA	Genetic Algorithm
FEB	Free Energy of Binding
HA	Hemagglutinin
HBA	Hydrogen Bond Acceptors
HBD	Hydrogen Bond Donors
ITC	Isothermal Titration Calorimetry
kb	Kilobase
kDa	KiloDalton
L.O.F	Lack of Fit
M2	Membrane2
MES	2-(<i>N</i> -morpholino)ethanesulfonic acid
MLR	Multiple Linear Regression
MTase	Methyl Transferase
NCI	National Cancer Institute
Neu5Ac	N-Acetylneuraminic Acid
NS	Non Structural

NTPase	N Triphosphatase
P	Peptide
PDB	Protein Data Bank
pred	Predicted
prM	preMembrane
PTC	Phase Transfer Catalyst
RFU	Relative Fluorescence Unit
RMSD	Root Mean Square Deviation
RTPase	RNA Triphosphatase
SA	Sialic Acid
sNS	Secreted Non Structural
S.O.R	Significant of Regression
TBAI	Tetrabutylammonium Iodide
WNV	West Nile Virus

REKABENTUK SECARA RASIONAL DAN SINTESIS PERENCAT UNTUK ENZIM NEURAMINIDASE H1N1 DAN PROTEASE DENGGI

ABSTRAK

Influenza dan denggi adalah dua daripada penyakit berjangkit yang disebabkan oleh virus. Rintangan virus terhadap ubat influenza komersial dan ketiadaan ubat untuk merencat virus denggi menggalakkan usaha untuk mencari perencat virus yang berpotensi. Setakat ini, enzim neuraminidase H1N1 ialah satu daripada sasaran utama dalam pencarian perencat influenza A manakala enzim protease NS2B-NS3 DENV2 pula merupakan sasaran utama dalam penemuan ubat denggi. Kajian sebelum ini mendapati bahawa asid ferulik yang dipencilkan daripada kulit buah manggis dapat merencat aktiviti neuraminidase H1N1 dengan nilai $IC_{50} = 200 \mu M$. Struktur aromatiknyanya yang sederhana menarik perhatian untuk dikembangkan sebagai perencat neuraminidase H1N1. Sebanyak 20 terbitan asid ferulik telah pun direkabentuk dan disintesis. Kajian asai atas sebatian tersebut menunjukkan nilai IC_{50} daripada 50 sehingga $> 1000 \mu M$. Rekabentuk hubungan kuantitatif struktur dan aktiviti menggunakan kaedah *Multiple Linear Regression* dilaksanakan untuk menghasilkan model yang menentukan hubungan positif daripada penderma dan penerima ikatan hidrogen dengan keputusan statistik yang baik ($r^2 = 0.758$; r^2 (adj) = 1.185; *Least-squared error* = 0.189). Dua model lagi direkabentuk berasaskan kajian hubungan kuantitatif struktur-aktiviti ini dan ianya diramalkan mempunyai IC_{50} lebih rendah daripada sebatian sebelumnya. Kajian sebelum ini mengenai penskrinan virtual daripada pangkalan data Institut Kanser Nasional terhadap protease NS2B-NS3 DENV2 menyarankan thioguanine sebagai perencat protease ini. Langkah penyelidikan yang sama juga dilaksanakan kepada

protease ini, menunjukkan nilai IC_{50} daripada 28 sehingga $> 1000 \mu M$. Kajian hubungan kuantitatif struktur-aktiviti pula dilaksanakan menggunakan kaedah *Genetic Function Approximation* dengan *Linear* terpilih sebagai model terbaik berasaskan keputusan statistik ($r^2 = 0.921$; r^2 (adj) = 0.884; r^2 (pred) = 0.820; *RMS residual errors* = 0.258; Friedman L.O.F = 0.141 and S.O.R *p-value* = $1.919e^{-006}$). Model ini menentukan koefisien pembahagi sebagai deskriptor positif untuk rekabentuk seterusnya. Sebanyak empat ligan baru direkabentuk, disintesis dan ditentukan aktiviti terhadap protease denggi. IC_{50} yang diramalkan dan ditentukan melalui aktiviti *in vitro* adalah saling setuju. Kesimpulannya, kaedah pemodelan *in silico* dalam kajian ini telah berjaya menunjukkan konsep merekabentuk perencat enzim secara rasional.

RATIONAL DESIGN AND SYNTHESIS OF INHIBITORS FOR H1N1 NEURAMINIDASE AND DENGUE PROTEASE ENZYME

ABSTRACT

Influenza and dengue are two of infectious diseases which caused by viruses. The viral resistances towards commercial anti-influenza as well as no drug available to combat dengue infection have prompted the search for potential inhibitors. Currently, H1N1 neuraminidase is one of the major targets in searching for inhibitor of influenza A as well as DENV2 NS2B-NS3 protease in dengue drug discovery. In a previous study, ferulic acid from *G. mangostana* pericarps has been isolated and showed an inhibition against H1N1 neuraminidase *in vitro* with $IC_{50} = 200 \mu M$. Its simple aromatic structure was attractive to be developed as H1N1 neuraminidase inhibitor. Twenty ferulic acid derivatives were designed *in silico* and synthesised, respectively. The *in vitro* assay showed the inhibitory activity with IC_{50} from 50 to $>1000 \mu M$. The Quantitative Structure-Activity Relationship (QSAR) modelling using Multiple Linear Regression was then carried out to produce the model defining a positive correlation of number of hydrogen bond donor as well as hydrogen bond acceptor with a good statistical results ($r^2 = 0.758$; r^2 (adj) = 1.185; Least-squared error = 0.189). Two further models were designed based on this QSAR equation and they were predicted to have lower IC_{50} values. On the other hands, previous study on virtual screening of National Cancer Institute database against DENV2 NS2B-NS3 protease suggested thioguanine as a potential inhibitor for the enzyme. The same approach carried out on DENV2 protease inhibitor design demonstrated the inhibitors possessing IC_{50} in the range of 28 to $>1000 \mu M$. QSAR models were also generated using Genetic Function Approximation selecting a Linear model as the best QSAR equation upon statistical results ($r^2 = 0.921$; r^2 (adj) = 0.884; r^2 (pred) =

0.820; *RMS residual errors* = 0.258; Friedman L.O.F = 0.141 and S.O.R *p-value* = 1.919e^{-006}). The model defined partition coefficient as the positive descriptors for further design. Four more ligands were then modeled, synthesised and tested their activity *in vitro*. The results showed a good agreement between its predicted and experimental IC₅₀. In general, the *in silico* modelling used in this study successfully proved the concept of the rational enzyme inhibitor design.

CHAPTER ONE

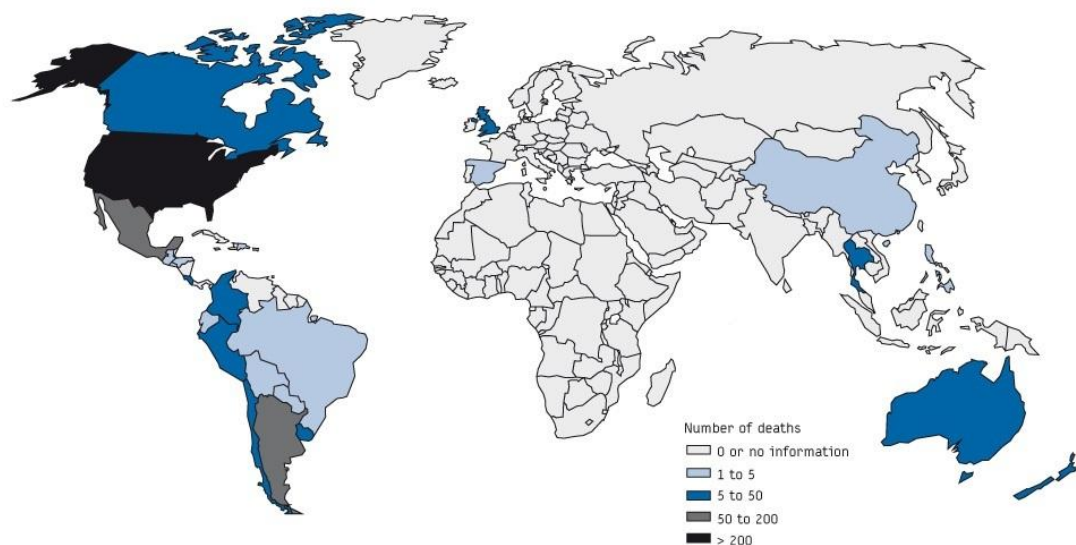
INTRODUCTION

1.1 Statement of the Problem

In recent years, world is threatened with the emergence of pandemic and endemic infections of viruses such as influenza A and dengue. The news of highly pathogenic influenza (H5N1) transmission from birds to human that resulted in 53 deaths in Vietnam, Cambodia and Thailand shocked the world in 2005 (WHO, 2005). More deaths were reported in the subsequent years and the threat of H5N1 now being compounded by the emergence of H1N1 pandemic in 2009. The World Health Organization (WHO) confirmed that the pandemic was spread to over 220 countries with more than 39 million cases and 15,417 deaths worldwide (see Figure 1.1) (CDC, 2010a; Fajardo-Dolci *et al.*, 2010). Compared to H5N1 avian influenza which emerged in 2005, H1N1 swine virus is less virulent but it is more prevalent than that of avian influenza (Salaam-Blyther, 2009).

FIGURE 1

Deaths associated with pandemic H1N1 influenza 2009 reported officially worldwide as of 16 July 2009



Source: Ministries of Health, local or national public health authorities, European Centre for Disease Prevention and Control, United States Centers for Disease Control and Prevention, World Health Organization.
Map drawn with Philcarto (free software available from: <http://philcarto.free.fr/>)

Figure 1.1 The death number of pandemic H1N1 influenza 2009 (CDC, 2010a).

Four years later in February 2013, a rare H7N9 subtype of avian influenza was isolated for the first time from the patient with pneumonia and acute respiratory distress syndrome in China. The infection showed alarming mortality rate (28% as of May 2013). Although H7N9 subtype is considered a low pathogen, the possibility of this subtype becomes resistant towards either vaccine or drug should be anticipated (Baranovich, 2013; Chen *et al.*, 2013).

Vaccines are available to prevent the infection of influenza. However, the existing vaccines have been mostly ineffective due to the emergence of rapidly variable mutation (Kim *et al.*, 1999; Zhang *et al.*, 2008a). Thus, the development of effective and safe anti-influenza becomes an urgent need (Gong *et al.*, 2007; Zhang *et al.*, 2008b).

Historically, the adamantane-based M2 ion channel protein inhibitors (amantadine and rimantadine) were the first drugs available for the treatment of influenza. Such drugs had only been useful in the treatment of Influenza A infection due to the fact that only the A strain of the virus has M2 ion channel protein (von Itzstein, 2007). The drugs inhibit the virus replication by blocking this ion channel via binding at the allosteric site which triggers a conformational change in the pore region. This action causes interfering proton transfer through the ion channel across the membrane of the virus or endosome. Therefore, the virus is unable to penetrate the host cell of membrane and the replication being stopped (Sandrock, 2010; Tisdale, 2009). Several toxic effects (CNS toxicity can manifest dizziness, nervousness and insomnia and also gastrointestinal effect such as nausea,

constipation and lost of appetite) have been reported along with rapid emergence of drug-resistant variants. Studies between 1994 and 2005 showed the increase of worldwide amantadine and rimantadine-resistance from 0.4% to 12.3%. Thus, the use of these drugs have been discouraged (Moscona, 2008).

Due to those disadvantages of M2 ion channel inhibitors, there was a shift to drug design targeting the other two surface glycoproteins; hemagglutinin (HA) and neuraminidase (NA), which are expressed by both influenza A and B. The HA plays as the receptor-binding and membrane fusion glycoprotein whilst the NA has a role as a receptor-destroying enzyme (An *et al.*, 2009). So far, no HA inhibitors have been clinically approved for influenza treatment (Meshram and Jungle, 2009).

Neuraminidase (NA), also known as sialidase, is the major surface glycoprotein that shows an important role in the viral replication, thus, has become an attractive target for anti-influenza drug. Zanamivir and oseltamivir are two examples of drugs that are effective for either A or B types of neuraminidase. Studies on NA active site and Structure-Activity Relationships (SARs) of published NA inhibitors disclose that the relative positions of the substituents (carboxylate, glycerol, acetamido, and hydroxyl) of the central ring mainly determine inhibition of the NA (Zhang *et al.*, 2008b).

Although zanamavir is highly effective, its inhalational delivery (D'Souza *et al.*, 2009; Hammad and Taha, 2009; Sun *et al.*, 2010) is not very attractive as oral delivery (via capsule/tablet) is much preferred. Oseltamivir overcomes this limitation, but the production cost is quite expensive as it relies on the expensive

starting material shikimic acid (Chand *et al.*, 1997). Although there have been a lot of efforts to discover new NA inhibitors with various scaffold, including aromatic (Chand *et al.*, 1997; Luo *et al.*, 1997), dihydropyran (Taylor *et al.*, 1998), cyclopentene (Babu *et al.*, 2009), cyclohexene (Lew *et al.*, 2000) and pyrrolidine (Zhang *et al.*, 2008b), the currently circulating clinical H274Y H1N1 mutant is quite resistant to oseltamivir (Yen *et al.*, 2006). Therefore, there is a need for new drugs that are cheaper and more effective against the influenza virus.

Next to another infectious disease, dengue has become a major disease burden for the South East Asian countries. Although the scale of dengue infection has not reached pandemic proportion, WHO currently estimates that there may be 50 million dengue infections worldwide every year (Murray *et al.*, 2013), with the infection endemic in South East Asian Regions. During 2008, in Indonesia, there were 17,604 cases reported and 10,000 to 12,000 of them concentrated over East and Central Java (WHO, 2014a). Furthermore, it was reported that in India, Indonesia, Sri Lanka and Thailand, the epidemic season started up to 4 weeks earlier than normal, in part due to the heavy rainy season in 2010. In the same year, Vietnam has occurrence rates 10 times greater than in 2009 (CDC, 2010). In Malaysia itself, this deadly disease had killed 107 people in 2005 and 102 people in 2004 (Chen *et al.*, 2006). In 2014, Bernama reported that dengue fever struck 17 deaths and 9,453 cases during January up to 2 February 2014, an increase over 100% of previous year (Bernama, 2014). It was reported that the case increased up close to 90,000 cases in December 2014. Figure 1.2 shows the dengue cases statistic in Malaysia from 1995-2014 ("Development of dengue vaccine: status sreview and future considerations," 2014). As of February 2015, the cases reached up to 59% more as compared to the

previous year, with 15,039 cases and 44 deaths as reported by Malaysian Department of Health (WHO, 2015).

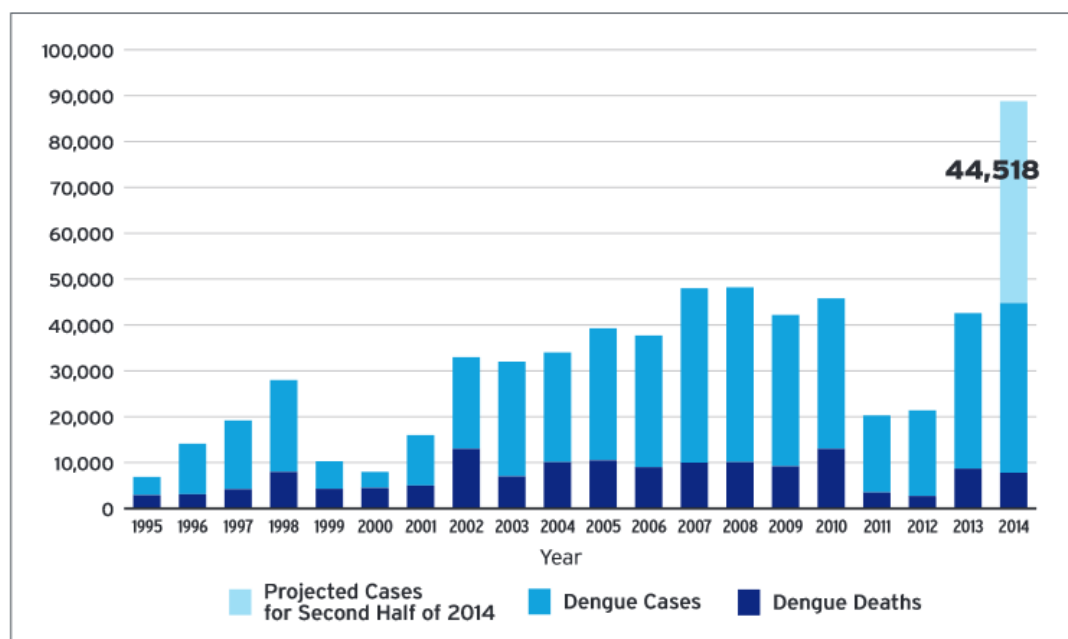


Figure 1.2 The dengue case statistic in Malaysia from 1995-2014 ("Development of dengue vaccine: status sreview and future considerations," 2014).

Dengue virus carries a positive single strand RNA in its genome. Serologically, this virus is divided into four serotypes, i.e. DENV1, DENV2, DENV3 and DENV4 (Jitendra and Vinay, 2011). Among those serotypes, DENV2 is the most prevalent type in dengue epidemic, especially in the South East Asian region. The virus genome is encoded by three structural proteins (C, prM, E) as well as seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) (Chambers *et al.*, 1990; Henschal and Putnak, 1990). Understanding of the virus life cycle promotes the clearer key targets of gene/protein in the replication of the virus, thus, is important in a drug design study. Currently, the serine protease of dengue virus has been a major target in dengue drug discovery (Luo *et al.*, 2008; Mueller *et al.*, 2008).

Presently, available vaccines e.g. live recombinant, DNA and subunit vaccine are limited by their ability to protect the body system from the viral infection on one serotype only, e.g. if the vaccine is sensitive against DENV1 only, thus the patient is still exposed to the dangers of being infected by other serotypes (DENV2-4). Therefore, the discovery of antiviral drug is of utmost concern in the treatment of dengue and its related diseases (Nair *et al.*, 2009; Tambunan *et al.*, 2011). Up to now, there is no established antiviral drug able to inhibit the dengue virus replication. There are only symptomatic treatment administered to the dengue patient to relieve the symptoms such as analgesic-antipyretic for fever and blood transfusion to recover the thrombolytic level in hemorrhagic patient (Kumar *et al.*, 2012; Zhou *et al.*, 2008). One of antiviral drugs, ribavirin, successfully inhibited the virus replication *in vitro* but it is limited in *in vivo* study due to the fact that the sugar moiety of this compound contributes to its poor bioavailability during oral administration (Sampath and Padmanabhan, 2009; Takhampunya *et al.*, 2006).

NS3 protease (NS3pro) is a trypsin like serine protease which plays a role in a post-translation from the genome to its proteins as well as its maturation. This enzyme has a catalytic triad made up by His51, Asp75 and Ser135 and enhanced by another non-structural protein named NS2B as an enzyme cofactor (Frimayanti *et al.*, 2011/2012; Yin *et al.*, 2006b; Yin *et al.*, 2006a). This cofactor activity is due to its hydrophilic region which is responsible for holding and promoting the activation of NS3 while the hydrophobic part playing around the membrane association upon the cleavage process (Chanprapaph *et al.*, 2005; Kee *et al.*, 2007; Niyomrattanakit *et al.*, 2004; Steuer *et al.*, 2009).

In most cases, the discovery of dengue antiviral by targeting NS2B-NS3pro activity was based on the non-prime substrates which were identified by profiling dengue virus using tetrapeptides (Freder and Miertus, 2010; Yang *et al.*, 2011). This peptidomimetic compound is able to inhibit the enzyme activity in a nanomolar concentration; however it is devoid of drug-like structure which can create a problem in physicochemical stability either during pharmaceutical preparation or in further pharmacokinetic step. As such, no peptidomimetic compound has been clinically used as dengue antiviral (Katzenmeier, 2004; Steuer *et al.*, 2011).

Within the last decade, the discovery of NS2B-NS3pro inhibitor have been made possible by small molecules design (Kiat *et al.*, 2006; Sidique *et al.*, 2009; Tomlinson and Watowich, 2011), assisted by Computer Aided Drug Design (CADD). These virtual experiments are very helpful to avoid the trials and errors when the tested compound examined its activity *in vitro* as well as *in vivo* (Acharya *et al.*, 2011; Korb *et al.*, 2009; Tang and Marshall, 2011). There are various methods such as structure based drug design (docking and molecular dynamics) and ligand based drug designs (pharmacophore and quantitative structure-activity relationship (QSAR)).

To date, there is no effective vaccine or antiviral drug available to protect against dengue diseases (Nair *et al.*, 2009; Tambunan *et al.*, 2011). Thus, as mortality and economic burden of this disease are quite high, major effort must be put in developing effective antiviral for the treatment against dengue.

Recently, a current unpublished study from Pharmaceutical Design and Simulation Laboratory, School of Pharmaceutical Sciences, USM, found that ferulic acid which was isolated from *Garcinia mangostana* pericarps showed a reasonable inhibition toward H1N1 neuraminidase with $IC_{50} = 200 \mu M$. Although this activity is still far away from the commercial drug (oseltamivir, $K_i = 3.78 \text{ nM}$) but ferulic acid structure can be used as a scaffold for NA inhibitor. Looking at the structure of ferulic acid, there are three functional groups identified probably contributing to H1N1 neuraminidase inhibition, i.e. carboxylic acid, hydroxyl, and methoxy groups. Furthermore, the ring system of aromatic compound is more planar than shikimic acid derivatives which showed bound to the NA active site and had high anti-neuraminidase activity (Chand *et al.*, 1997). Ferulic acid has a highly correlated structure with vanillin (see Figure 1.3). Ferulic acid can be prepared synthetically by reacting vanillin and malonic acid; and vice versa, vanillin can be produced by hydrolyzing ferulic acid at a certain condition.

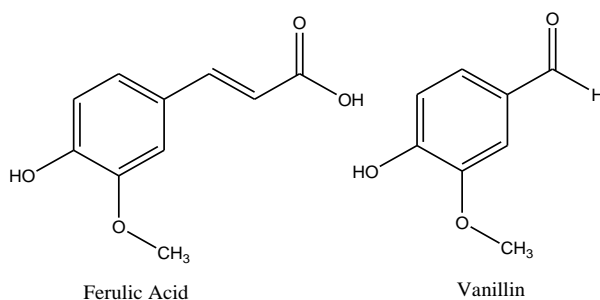


Figure 1.3 The structure of ferulic acid and vanillin.

On the other hand, also unpublished previous works from Pharmaceutical Design and Simulation (PhDs) laboratory, USM, found four National Cancer Institute (NCI) compounds possessed bioactivity *in vitro* against DENV2 NS2B-NS3pro from their virtual screening. Those four NCI compounds showed some hydrogen bond interactions to amino acid residues which are important for DENV2

NS2B-NS3pro activities. The docking results corresponded well with the *in vitro* study. These four compounds, however are not typically classified as peptidomimetic, therefore, there is an opportunity to develop small molecules to be the lead compound of dengue antivirus. Among those NCI compounds, the diversity0713 (see Figure 1.4) which has a thioguanine scaffold is considered chemically accessible in term of synthetic route as well as the availability of starting materials.

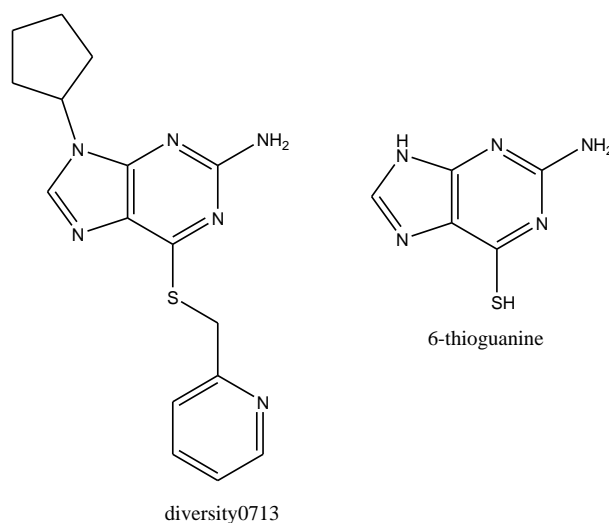


Figure 1.4 The structure of diversity0713 (www.pubchem.com) and thioguanine.

1.2 Objectives

The goal of this study is to discover novel neuraminidase and dengue antivirus.

Specifically the objectives are:

1. To model and design a series of H1N1 NA and DENV2 NS2B-NS3pro inhibitors bearing ferulic acid and thioguanine scaffold respectively, using *in silico* methods.
2. To synthesise the designed H1N1 NA and DENV2 NS2B-NS3pro inhibitors.

3. To evaluate the *in vitro* activities of the synthetic compounds (H1N1 NA and DENV2 NS2B-NS3pro inhibitors) against H1N1 NA and DENV2 NS2B-NS3pro, respectively.
4. To study the quantitative structure-activity relationship (QSAR) of the synthetic compounds, ferulic acid and thioguanine derivatives as the H1N1 NA and DENV2 NS2B-NS3pro inhibitors, respectively.

CHAPTER TWO

LITERATURE REVIEW

2.1 Influenza A and Its Drug Treatment

2.1.1 Influenza A virus

Influenza A virus is a pathogen which commonly infects the upper respiratory tract causes several symptoms such as myalgia, malaise, and fever for a few days as common respiratory symptoms. The condition can be more complicated when it affects other cells or organs causes pneumonia and myocardial infection (Kuszewski and Brydak, 2000; Svetlikova *et al.*, 2010). The major pandemics were reported in three periods during 20th centuries: 1918-1921 (H1N1”Spanish” Influenza), 1957-1958 (H2N2”Asian” Influenza) and 1968 (H3N2”Hongkong”Influenza) (Gautret *et al.*, 2010). Since 1997 to 2005, H5N1 influenza became dominant in virus circulation among birds but later on, it kills hundreds people worldwide (Saif and Espinoza, 2006). The swine flu (H1N1) virus was firstly found having human to human transmission in Mexico (2009) but luckily, this virus is less pathogenic than H5N1 virus (Simonsen *et al.*, 2011). The updated status of H1N1 reported by WHO in 2009 indicated this influenza virus found in 381 human specimens in Washington DC (WHO, 2014a).

Influenza A virus is a negative-strand RNA virus belonging to orthomyxo viruses. The shape of the virus is a pleomorphic particle, typically having a spherical or long ‘lollypop’ like filament shape. The enveloped-negative RNA segmented genomes are packed into a nucleocapsid which complexes with protein polymerase. This RNA-protein (RNPs) complex is packed in a lipoprotein envelope lined by three

surface proteins: hemagglutinin (HA), neuraminidase (NA), and the M2 proton ion channel (Magano, 2009; Shtyrya *et al.*, 2009).

Figure 2.1 shows the schematic diagramme of Influenza A virus. Hemagglutinin (HA) is a glycoprotein which recognizes the cell targets by binding to sialic acid receptor on the host cell membrane. This protein binds to the receptor via α -ketosidic-linked terminal that are followed by fusion of the viral through the endosomal membranes succeeding endocytosis. After endocytosis, the pH of the cell changes and triggers the prolongation of the central coiled at the N-terminus thus drives out the fusion peptide and penetrates into a cellular membrane (Isin *et al.*, 2002; Sollner, 2004; von Itzstein, 2007). There are 15 subtypes of HA (H1-H15); three of them: H1, H2 and H3 attack humans by binding at the sialic acid receptor in the respiratory tract; while another subtype, H5, invades protein in avian digestive enzyme (Fouchier *et al.*, 2005). The swine is susceptible toward both human and avian influenza virus, thus the novel strain of H1N1 can be generated by those two virus reassortment in this species leading to the theory called as “mixing vessel” (Ma *et al.*, 2007).

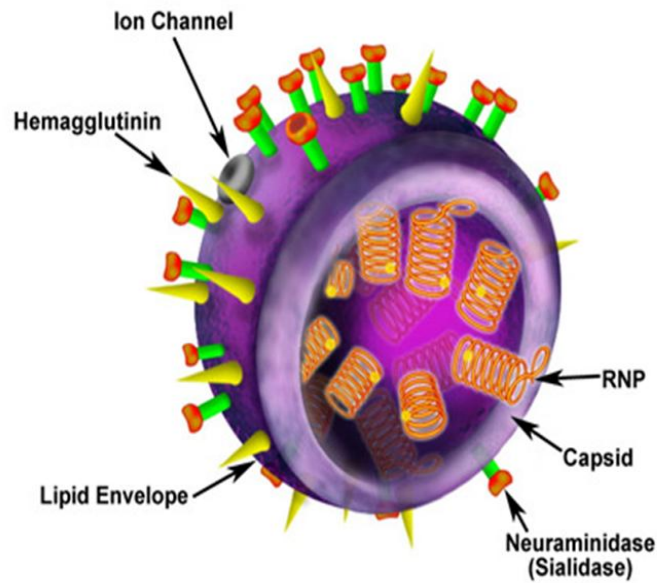


Figure 2.1 The influenza viral lipid envelope with a nucleocapsid containing three surface proteins: hemagglutinin (HA), neuraminidase (NA), and the M2 proton ion channel. The genome of the viral RNAs are presented as red coils bound to Ribonuclear Proteins (RNPs) (Betakova, 2007).

Neuraminidase (NA), also known as sialidase, exists as a tetramer of identical subunits. It is normally attached to the virus surface through a long protein stalk. The active sites are in a deep depression on the upper surface. They bind to polysaccharide chains of the sialic acid receptor and clip off the sugars (sialic acid) at the terminal end. The surface of neuraminidase is decorated with several polysaccharide chains (seen extending upwards and downwards in this structure, Figure 1.1) that are similar to the polysaccharide chains that decorate our own cell surface proteins. NA can be divided into 9 subtypes (N1-N9) and it is important for viral replication and infection (Du *et al.*, 2007; Tisoncik *et al.*, 2011). This enzyme cleaves the terminal sialic acid moieties from the receptors to facilitate release of the virion progeny from the infected cell. NA is able to facilitate the early processing of influenza virus infection in the lung epithelial cells. Due to these essential roles, NA

has been an attractive target in anti-influenza discovery (Du *et al.*, 2007; Gong *et al.*, 2007; Platis *et al.*, 2006; Wang *et al.*, 2010).

M2 ion protein channel is the third membrane protein which provides the virus structural integrity by permitting the proton to enter the virus particle during un-coating of the virion in the endosome (Bauer *et al.*, 1999). It is a homotetramer consisting of four polypeptide chains from 96 amino acids, with the structural domains: an amino-terminal extracellular domain (comprising 23 residues), as a single internal hydrophobic domain that acts as a trans-membrane domain (19 residues) and 54-residue cytoplasmic tails. This membrane protein is important to prevent inactivation of progeny virus as well as premature acid activation of the newly synthesised HA (Betakova, 2007; Rossman *et al.*, 2010).

2.1.2 Influenza A neuraminidase

The crystal structure of neuraminidase first solved in 1983 (Varghese *et al.*, 1983) has given a glimpse on the mechanism of action of this enzyme at a molecular level. This structure was of N2 subtype at a resolution of 2.9 Å, showing that the enzyme adopts a tetramer shape (see Figure 2.2) with the total molecular weight of 240 kDa. Each monomer is formed by six 4-stranded of anti-parallel β -sheets which is arranged as blade propellers around the central of pseudo six fold. The first strand of each sheet is parallel to the central of the propeller and the outer strand is vertical toward it, causes each strand being twisted. The outermost strand of the first sheet is linked to the central of strand for the next sheets. Loops connecting these strands contain a diverse important amino acids and form enzymatic site on its upper surface.



Figure 2.2 The tetramer shapes of neuraminidase N2 (Varghese *et al.*, 1983).

Phylogenetic studies have divided the neuraminidase influenza virus into two groups, Group 1 contains subtypes N1, N4, N5 and N8 and Group 2, subtypes N2, N3, N6, N7 and N9 (Gong, 2007; Tisdale, 2009). The main difference between these two groups is the presence of a loop consisting amino acids 147 to 151 which is also named as 150-loop (Xu *et al.*, 2008). For example, the 150-loop of the N1 has Val149 while for N9 it is formed by Ile149. For both groups, the 150-loop contains an amino acid Asp151 as shown in the superposition of N1 and N9 in Figure 2.3. However, in N1, its side chain has a carboxylic acid steering away from the active site of the enzyme in an open conformation, while in N9, it is a close form (Du *et al.*, 2007).

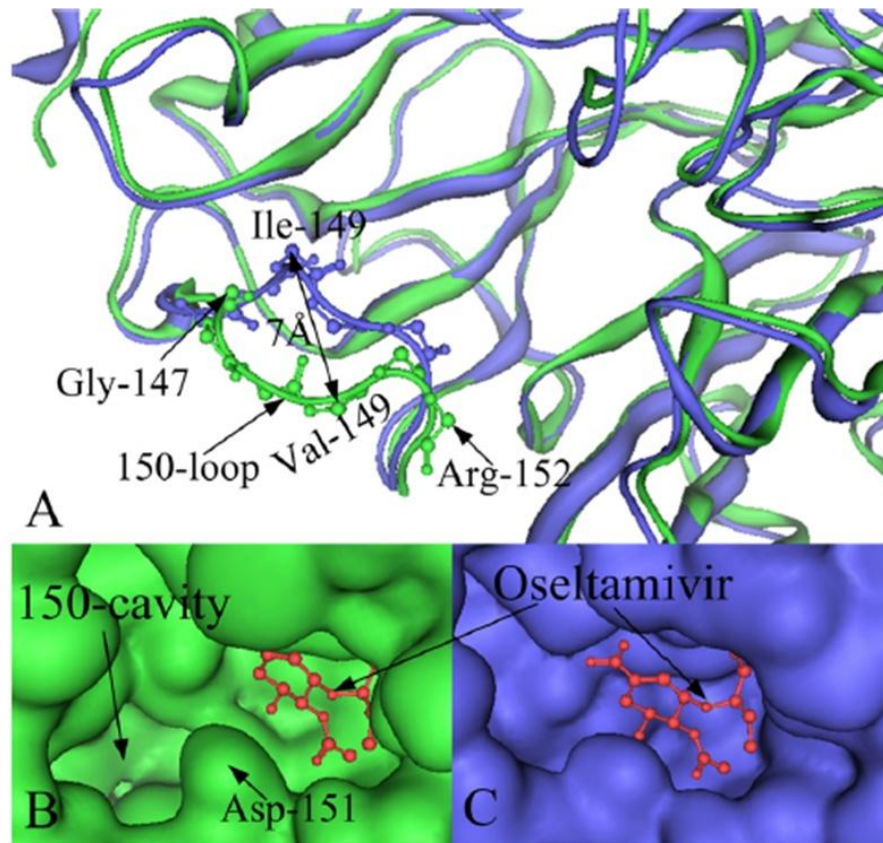


Figure 2.3 The comparison of crystal structures between N1 and N9. (A) A superposition of N1 (PDBID 2HTY) and N9 enzyme (PDBID 1F8B) where the key residues are shown in ball-and-stick forms. (B) Ligand in the active cavity of 2HTY with adjacent 150-loop. (C) Ligand in the active cavity of 1F8B (Du *et al.*, 2007).

The function of NA is to facilitate mobility of the virus, both to and from the site of infection. NA catalyzes the cleavages of (2-6) - or α (2-3)-ketosidic linkage between a terminal SA and inward-face sugar residue. This broken bond facilitates to spread the virus in the respiratory tract and allows the elution of progeny virus from the infected cells. The removal of sialic acid from the oligosaccharide moiety of HA and NA also helps to prevent the virus self-aggregation after leaving the host cells (Gong *et al.*, 2007). The mechanism of neuraminidase activity that facilitates the viral replication is shown in Figure 2.4.

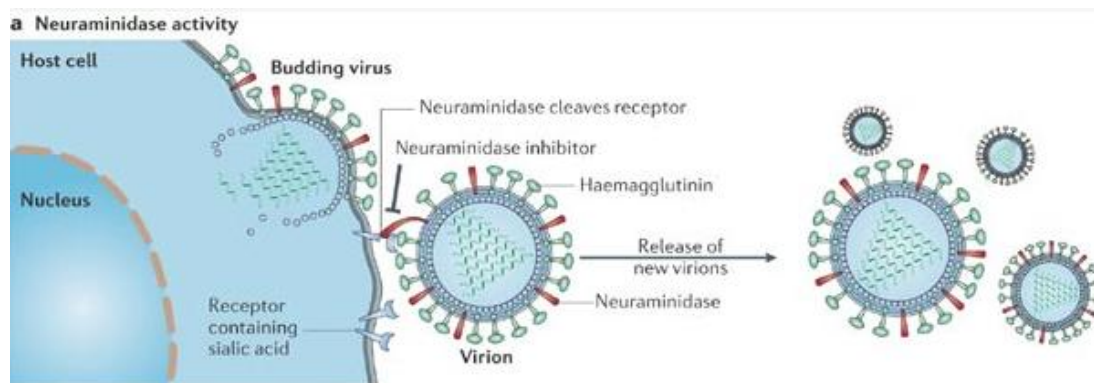


Figure 2.4 The cleavage of the new synthesised virus from its sialic acid receptor by neuraminidase (Clercq, 2006).

The availability of X-ray crystal structures of neuraminidase with high-resolutions complexed with sialic acid (SA, N-acetylneuraminic acid, Neu5Ac) has improved the understanding of the action mechanism of this enzyme. Figure 2.5 described the active binding site of neuraminidase when co-crystallized with Neu5Ac. This active site is highly conserved and presents a rigid catalytic center (Gong *et al.*, 2007; von Itzstein, 2007; Yen *et al.*, 2006). There are eight amino acids which directly make contact with Neu5Ac which provide conserved binding by the charge-charge interaction between the carboxylate group and three positively charged arginine residues (Arg118, Arg292 and Arg371). The NH group of 5-N-Acetyl interacts with the active site cavity via hydrogen bonding with water molecule while the oxygen carbonyl on the same 5-N-Acetyl moiety of Neu5Ac bonds to N of Arg152 via direct hydrogen bonding while its two hydroxyl groups of glycerol side chain are bonded to the carboxylate oxygens of Glu276. In addition, the 2-hydroxyl of Neu5Ac makes a direct contact with the carboxylate oxygen of Asp151 (Varghese, 1999; von Itzstein, 2007).

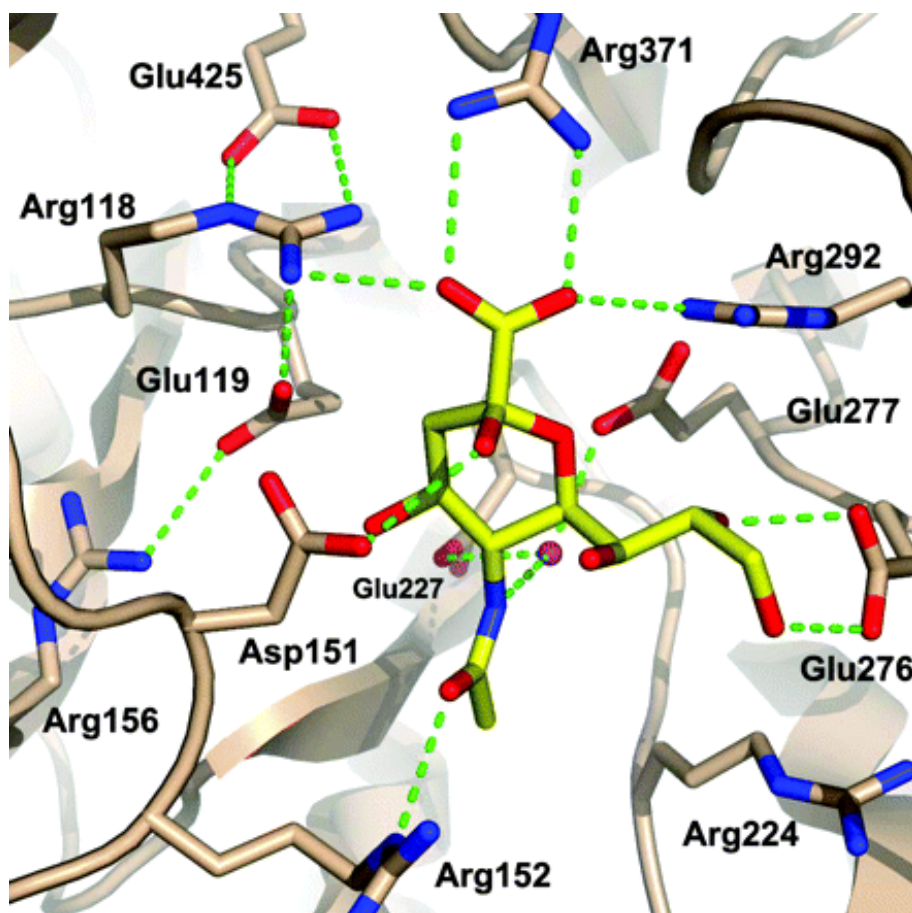


Figure 2.5 The binding site of A/Tokyo/3/67 (H2N2) influenza virus NA and sialic acid (PDBID 2BAT). The green dot lines describe the hydrogen bonding between ligand and specific amino acid residues of NA's active site (Yen *et al.*, 2006).

2.1.3 Neuraminidase Inhibitors

Most potent NA inhibitors were developed based on the structural information of the N2 NA conserved active site and its complex with sialic acid. Figure 2.6 shows the two-dimensional 'airplane' model to illustrate the NA active site (Wang *et al.*, 2010). This 'airplane' model was used to define the Structure-Activity Relationships (SARs) of NA inhibitors, which is mainly determined by the relative positions of the substituents (carboxylate, glycerol, acetamido and hydroxyl) in the central ring of the inhibitor. There are four well conserved active site domains which consist of Site 1 (Arg118, Arg292, Arg371), Site 2 (Asp151, Glu119,

Glu227), Site 3 (Ile222, Trp178) and Site 4 (Glu276, Glu277) (Gong *et al.*, 2007; Zhang *et al.*, 2008a).

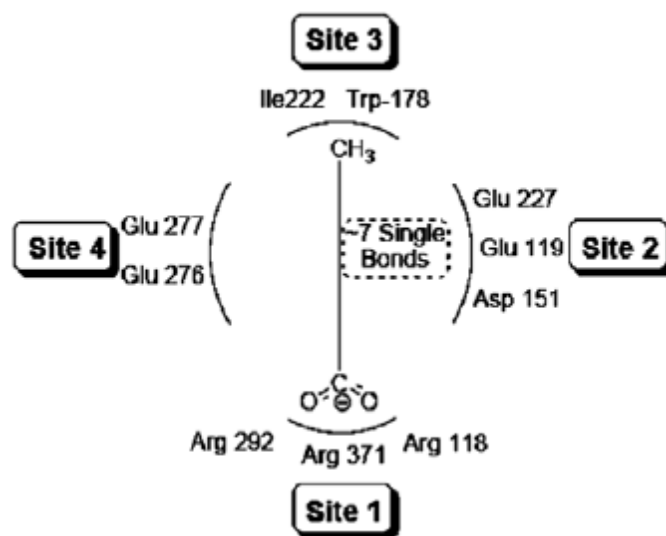


Figure 2.6 ‘Airplane’ model of NA active site (Zhang *et al.*, 2008a).

The elucidation of the three-dimensional structure of influenza neuraminidase located the catalytic site, which makes the design of highly effective inhibitors became feasible. The first NA inhibitor is 2-deoxy-2,3-didehydro-*N*-acetylneuraminic acid (DANA (**1**), Figure 2.7) synthesised in 1969 (Chen *et al.*, 2013). It was designed as a mechanism-based inhibitor and was proposed to be a transition state analogue with $K_i \sim 1\text{pM}$. Compound **1** showed a good activity *in vitro* but was found inactive as antivirals in an animal model (Laver and Elspeth, 2002). Earlier structure based drug design work, showed the availability of space between **1** and the enzyme in the vicinity of the 4-hydroxyl binding pocket, and thus, it was desirable to fill that space with basic substituents on the sugar (Smith *et al.*, 2001). Upon this information, the compounds targeted for synthesis were 4-amino-Neu5Ac2en (**2**) and 4-guanidino-Neu5Ac2en (**3**) (Figure 2.7). The K_i of these compounds are 20 and 5000 times respectively improved over that for **1** when

assayed against the neuraminidase of the viral isolate from which the structure had been determined.

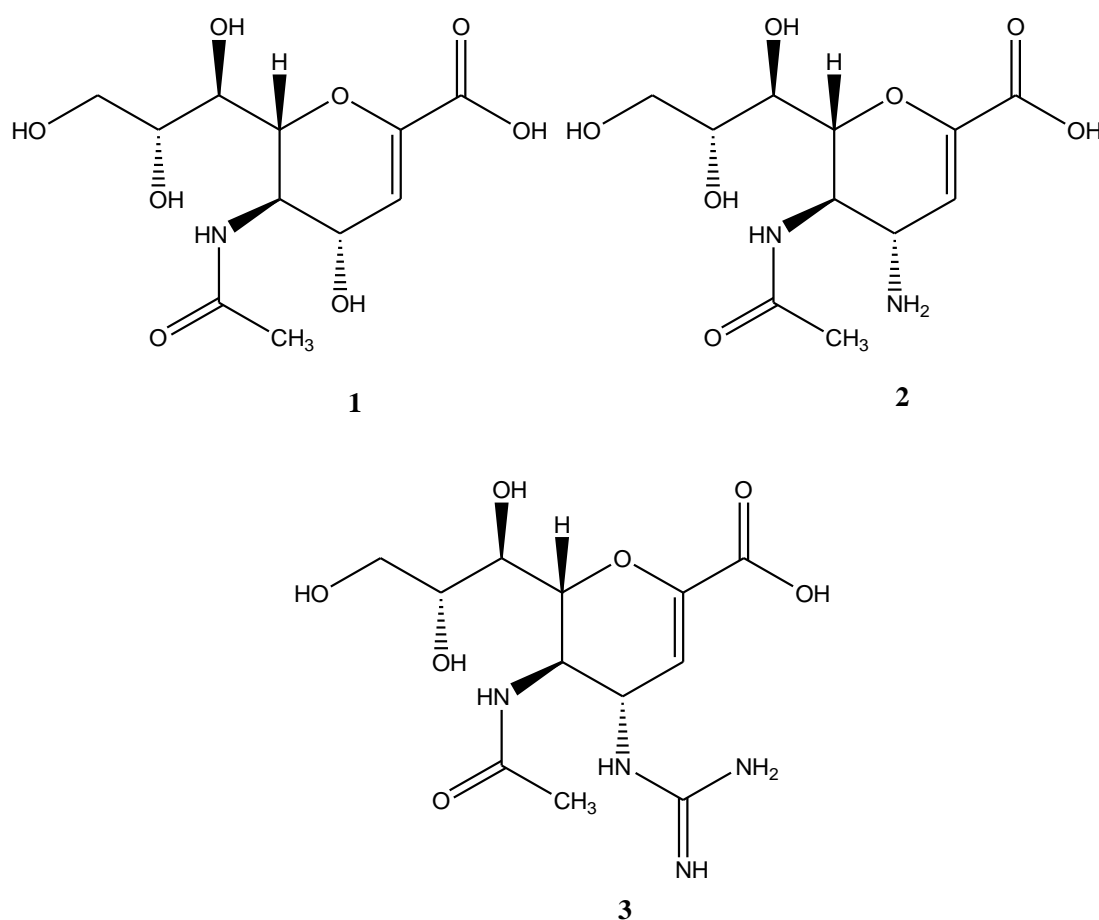


Figure 2.7 The structure of NA inhibitors from sialic acid derivatives (Tapar *et al.*, 2011).

Compound **3** or its drug name Zanamivir has an oral inhalation administration not only due to its high polarity, but also consciously intended for direct delivery to the respiratory tract, the principal site of the viral replication (Gupta *et al.*, 2011; Sun *et al.*, 2010). This sialic acid derivative showed a potent antiviral activity *in vitro* against influenza A as well as B viruses, including amantadine- and rimantadine-resistant isolates. Further study also showed that **3** is active against avian influenza A viruses, including influenza A H5N1, H6N1, H7N7 and H9N2 as well as some influenza strains resistant to oseltamivir (Elliot, 2001).

A new orally active neuraminidase inhibitor, coded by GS4071 (**4**) was designed to overcome the problem of oral bioavailability of zanamivir (Zhang *et al.*, 1997). The decrease in polarity contributed by carboxyl, hydroxyl as well as guanidine in zanamivir was achieved in **4** by removing the heterocyclic oxygen but augmenting the isopentyl as shown in Figure 2.8.

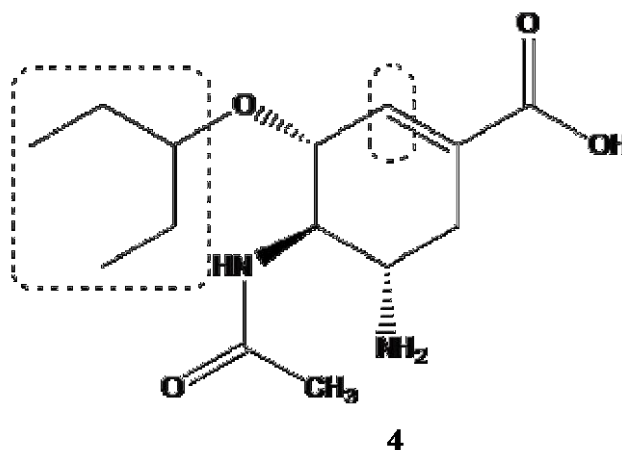


Figure 2.8 The structure of **4** with the dot boxes describes the nonpolar site of the particular structure.

The similarity of this compound to the compound **2** is evident especially their binding mode as demonstrated from the X-ray crystal structure of the complexes bound to neuraminidase (Figure 2.9). The carboxylate, acetyl as well as amino group belongs to the compound **4** were shown to bind to the conserved amino acid residues of neuraminidase's active site. In addition, the isopentyl group makes several favorable hydrophobic contacts with amino acid residues Ile222, Arg224, and Ala246, increasing the binding affinity (Lew *et al.*, 2000).

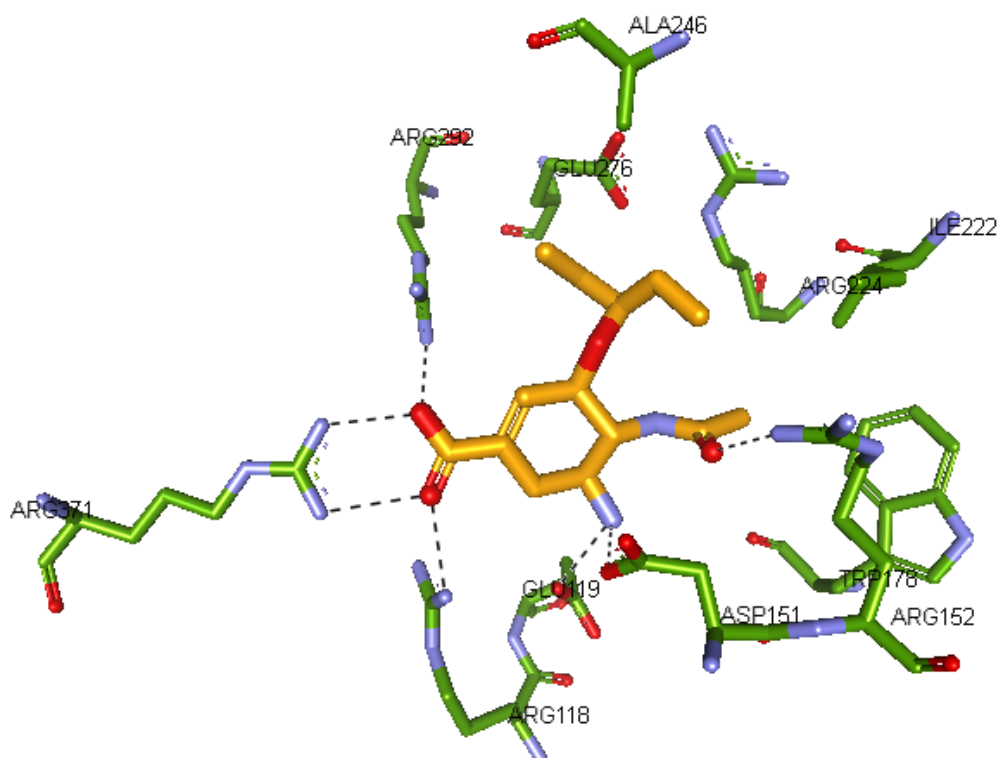


Figure 2.9 The X-ray crystal structure of neuraminidase complex with **4**. The hydrophobic interactions are indicated by the closeness between isopentyl group and amino acid residues Ile222, Arg224 and Ala246 (Lew *et al.*, 2000).

In pharmacokinetic experiments, **4** demonstrated only ~5% bioavailability in rats which is similar to **3**. Somehow, the conversion of carboxylic acid to its ester form, oseltamivir, provides five folds higher bioavailability than **4** when it is administered orally. It was found that the ethyl ester form of this compound metabolized via hydrolysis reaction to become carboxylate, the active metabolite which is shown to have poor bioavailability. Therefore, the phosphate salt of oseltamivir was developed to overcome the bioavailability problem thus, the compound can be administered orally (D'Souza *et al.*, 2009; Sun *et al.*, 2010).

The recent X-ray crystallographic study reveals that influenza A virus of the group-1 NA (N1) contains a larger cavity adjacent to the active site, formed by

residue 147-152 (150-loop) which is not found in the group 2 NAs. The group 1 NAs can bind ligands in the open as well as in the closed conformation. However, only the close conformational state was found in the group 2 NAs. The flexible 150-loop residues are occupied on the vicinity by the ligand and interact mostly with the amino and acetamido groups of **4**. The active sites are still conserved at the triad arginines (Arg118, Arg292, and Arg371) that interact with the carboxylate group. Moreover, the affinity is increased by the interaction between the acetamido group and Glu276 that forms hydrogen bonds with the substrate of hydroxyl groups (Collins *et al.*, 2008; Rungrotmongkol *et al.*, 2009).

A new scaffold containing five members of alicyclic compound was introduced to generate the antiviral activity against influenza virus via neuraminidase inhibition (Bianco *et al.*, 2005). Peramivir (**5**) with the cyclopentane core attached with carboxylic acid as well as acetlamino group and showed to possess *in vitro* activity against H1N1 virus with $IC_{50} = 0.38 \mu M$ (Ikematsu *et al.*, 2012). The crystal structure of N8 complex with Peramivir (PDBID 2HTU) shows the conserved binding mode of this compound to the enzyme (see Figure 2.10).

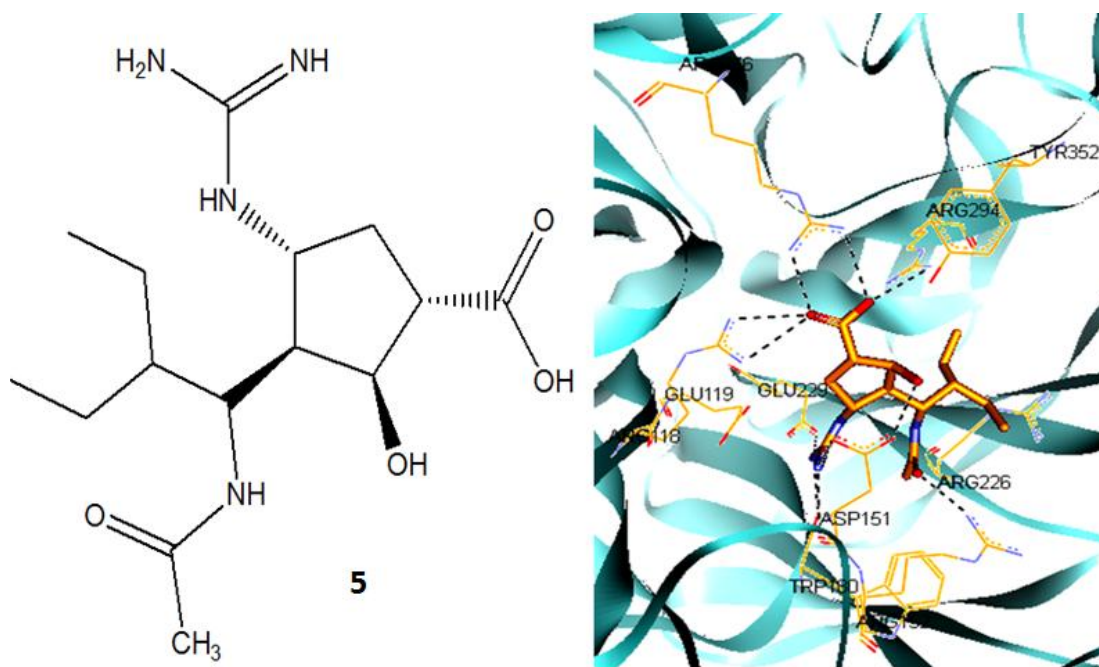


Figure 2.10 The structure of **5** in 2D as well as its 3D structure complex to N8 NA (PDBID 2HTU) was generated using Discovery Studio 2.5.

The latest NA inhibitor clinically approved in Japan in 2010 is Laninamivir which is proven effective against oseltamivir-resistant mutant. This drug is also prepared in octanoate ester form in the major structure of **3**. This laninamivir ester (**6**) is unique in its ability to rotate the Glu276 to form a salt bridge with Arg224, thus giving more space for Asn294 to interact with the the 9-ester-O of Laninamivir when it was crystallized with 2009 H1N1 NA (p09N1) (see Figure 2.11) (Ikematsu and Kawai, 2011; Vavricka *et al.*, 2011).